EVALUATION OF NPDES PHASE 1 MUNICIPAL STORMWATER MONITORING DATA

Robert Pitt, Alex Maestre, and Renee Morquecho
Dept. of Civil and Environmental Engineering
University of Alabama
Tuscaloosa, AL 35487

Ted Brown, Chris Swann, Karen Cappiella, and Tom Schueler
Center for Watershed Protection
Ellicott City, Maryland 21043

Abstract

The University of Alabama and the Center for Watershed Protection were awarded an EPA Office of Water 104(b)3 grant in 2001 to collect and evaluate stormwater data from a representative number of NPDES (National Pollutant Discharge Elimination System) MS4 (municipal separate storm sewer system) municipal stormwater permit holders. The data are being collected and reviewed to both describe the characteristics of this data and to provide guidance to permit writers for future sampling needs.

There have been serious concerns about the reliability and utility of Phase 1 stormwater NPDES monitoring data, mainly due to the wide variety of experimental designs, sampling procedures, and analytical techniques used. On the other hand, the cumulative value of the monitoring data collected over nearly a ten year period from more than 200 municipalities throughout the country has a great potential in characterizing the quality of stormwater runoff and comparing it against historical benchmarks. This project is creating a national database of Phase 1 stormwater monitoring data, providing a scientific analysis of the data, and providing recommendations for improving the quality and management value of future NPDES monitoring efforts.

Each data set is receiving a quality assurance/quality control review, based on reasonableness of data, extreme values, relationships among parameters, sampling methods, and a review of the analytical methods. The statistical analyses is being conducted at several levels. Probability plots are used to identify range, randomness and normality. Clustering and principal component analyses are also being utilized to characterize significant factors affecting the data patterns. The master data set is also being evaluated to develop descriptive statistics, such as measures of central tendency and standard errors. We are testing for regional and climatic differences, the influences of land use, and the effects of storm size and season, among other factors.

This paper describes our data collected to date and presents some preliminary data summaries. We have been collecting much data to date, and encourage any other communities with wet weather outfall data collected as part of their NPDES permit program to contact us so we can include as much data as possible in our final effort.

Project Description and Background

The importance of this project is based on the scarcity of nationally summarized and accessible data from the existing NPDES stormwater permit program. There have been some local and regional data summaries, but little has been done with nationwide data. A notable exception is the CDM national stormwater database (Smullen and Cave 2002) that combined historical NURP (Nationwide Urban Runoff Program) (EPA 1983), available urban USGS, and selected NPDES data. Their main effort has been to describe the probability distributions of this data (and corresponding EMCs, the event mean concentrations). They concluded that concentrations for different land uses were not significantly different, so all their data was pooled.

Other regional databases also exist, mostly using local NPDES data. These include the Los Angeles area database, the Santa Clara and Alameda County (CA) databases, the Oregon Association of Clean Water Agencies Database, and the Dallas area stormwater database. These regional data are (or will be) included in this comprehensive NPDES national database. However, we will not be including the USGS or historical NURP data in this NPDES database due to lack of consistent descriptive information for the older drainage areas. Much of the NURP data are available in electronic form at the University of Alabama student American Water Resources Association web page at: http://www.eng.ua.edu/~awra/download.htm The results from these other databases will be compared to our results during our final analyses to indicate any important differences.

This new NPDES database is unique in that detailed descriptions of the test areas and sampling conditions are also being collected, including aerial photographs and topographic maps for many locations which we are collecting from public domain Internet sources. The land use information used is as supplied by the communities submitting the data, although aerial photographs and maps are also used to clarify any questions. Most of the sites have homogeneous land uses, although many are mixed. These characteristics are all fully noted in the database.

This project is collecting stormwater runoff data from existing NPDES permit applications and permit monitoring reports; we are conducting QA/QC (quality assurance/quality control) evaluations of these data; and statistical analyses and summaries of these data. The final information will be published on the Internet (such as on an EPA OW-OWM, Office of Water and Office of Wastewater Management, site and on the Center for Watershed Protection's SMRC, Stormwater Manager's Resources Center, site at: http://www.stormwatercenter.net/). Some of the information is currently located at Pitt's teaching and research web site at: http://www.eng.ua.edu/~rpitt/.

The phase 1 NPDES communities included areas with:

- A stormwater discharge from a MS4 serving a population of 250,000 or more (large system), or
- A stormwater discharge from a MS4 serving a population of 100,000 or more, but less than 250,000 (medium system)

More than 200 municipalities, plus numerous additional special districts and governmental agencies were included in this program. Part 2 of the NPDES discharge permit application specified that sampling was needed and that the following was to be included in the application:

- Proposed monitoring program for representative data collection during the term of the permit.
- Quantitative data from 5 to 10 representative locations,

- Estimates of the annual pollutant load and event mean concentration (EMC) of system discharges,
- Proposed schedule to provide estimates of seasonal pollutant loads and the EMC for certain detected constituents during the term of the permit.

The permit applications were due in 1992 and 1993. For Part 2 of the application, municipalities were to submit grab (for certain pollutants) and flow-weighted sampling data from selected sites (5 to 10 outfalls) for 3 representative storm events at least 1 month apart. In addition, the municipalities must have also developed programs for future sampling activities that specified sampling locations, frequency, pollutants to be analyzed, and sampling equipment.

Numerous constituents were to be analyzed, including typical conventional pollutants (TSS, TDS, COD, BOD_5, oil and grease, fecal coliforms, fecal strep., pH, Cl, TKN, NO_3, TP, and PO_4), plus many heavy metals (including total forms of arsenic, chromium, copper, lead, mercury, and zinc, plus others), and numerous listed organic toxicants (including PAHs, pesticides, and PCBs). Many communities also analyzed samples for filtered forms of the heavy metals. Our database includes information for about 125 different stormwater quality constituents, although the current database is mostly populated with data from 44 of the commonly analyzed pollutants (as summarized later in Table 3). Therefore, there has been a substantial amount of data collected during the past 8 or 9 years from throughout the country, although most of these data are not readily available, nor have detailed statistical analyses been conducted and presented.

Data Collection and Analysis Efforts to Date

As of mid-December 2002, 3,757 events from 66 agencies and municipalities from 17 states have been collected and entered into our database. These locations are listed in Table 1. Table 2 lists 27 states where municipalities have been contacted and we plan to target for our next phase of data collection. Figure 1 shows the locations of these municipalities on a national map. We anticipate excellent national coverage, although we may have few municipalities from the northern west-central states of Montana, Wyoming, North and South Dakota (where cities are generally small, and few were included in the Phase 1 NPDES program).

Some of the municipalities that we have contacted (and some where we actually received data) have information that could not be used for various reasons. One of the most common reasons for not being able to use the data was that the samples had been collected from receiving waters (such as Washington state, Nashville, and Chattanooga). We are using data only from well-described stormwater outfall locations. These can be open channel outfalls in completely developed areas, but are more commonly conventional outfall pipes. The other major problem is that the sampling locations and/or the drainage areas were not described. We are using data with some missing information for now, with the intention of obtaining the needed information later. However, there will likely still be some minor data gaps that we will not be able to fill. In addition, the list of constituents being monitored has varied for different locations. Most areas evaluated the common stormwater constituents, but few have included organic toxicants. The most serious gap is the frequent lack of runoff volume data, although all sites have included rain data. Finally, if we collect all the data we have asked for, our current project resources will not permit us to fully utilize them, as it requires a great deal of time to enter and review this information.

The assembled data has been entered into a database which contains site descriptions (state, municipality, land use components, and EPA rain zone), sampling information (date, season, rain depth, runoff depth, sampling method, sample type, etc.), and constituent measurements (concentrations, grouped in categories).

In addition, more detailed site, sampling, and analysis information has been collected for each sampling site and included as supplemental information. We are using the reported land use information supplied by the communities, and are verifying some with aerial photographs and maps. In many cases, the sampled watersheds have multiple land uses and those designations are included in the database (we list the percentages of the drainage as residential, commercial, industrial, freeway, institutional, and open space). Our final data analyses will consider these mixed sites also, although the following preliminary results are only for the homogeneous land use sites.

Preliminary Summary of Phase 1 Stormwater Data

We plan to acquire additional stormwater data before our final data analysis, and to complete many of the missing records. The following data and analysis descriptions should be considered preliminary and will change with these additional data and analyses. However, we are presenting only our most basic and robust analyses here for consideration. Our final report and data presentations will obviously be much more comprehensive.

Table 1. Municipalities whose Data has been Entered into Database

ALABAMA
- Jefferson County
- Mobile

ARIZONA
- Maricopa County
- Tucson

CALIFORNIA
- Alameda
- Caltrans

COLORADO
- Denver
- Colorado Springs

GEORGIA
- Clayton County
- Cobb County
- De Kalb County
- Fulton County
- Gwinnett County
- Atlanta

IDAHO
- Ada County Highway District

KANSAS
- Topeka
- Wichita

KENTUCKY
- Jefferson County
- Louisville
- Lexington

MASSACHUSETTS
- Boston

MARYLAND
- Anne Arundel County
- Baltimore County
- Baltimore City
- Carroll County
- Charles County
- Harford County
- Howard County
- Montgomery County
- Prince Georges County
- State Highway

MINNESOTA
- Minneapolis

NORTH CAROLINA
- Charlotte
- Fayetteville
- Greensboro

OREGON
- Clackamas County
- Eugene
- Gresham
- Portland
- Salem
- ODOT

PENNSYLVANIA
- Philadelphia

TENNESSEE
- Knoxville
- Memphis

TEXAS
- Arlington
- Dallas
- Dallas County
- Fort Worth
- Garland
- Harris County
- Houston
- Irving
- Mesquite
- Plano
- Tarrant County

VIRGINIA
- Arlington County
- Chesapeake County
- Chesterfield County
- Fairfax County
- Hampton County
- Henrico County
- Newport News County
- Norfolk County
- Portsmouth County
- Virginia Beach County

Table 2. Communities Targeted for Next Phase of Data Collection

State	Communities
ALABAMA	Madison Huntsville - Madison Montgomery
ALASKA	Anchorage
ARIZONA	Pima County Mesa Phoenix Tempe
CALIFORNIA	Various Communities
COLORADO	Aurora Lakewood Littleton
DELAWARE	Wilmington New Castle County
FLORIDA	Various Communities
HAWAII	Honolulu County
ILLINOIS	Rockford
INDIANA	Indianapolis
KANSAS	Kansas City
LOUISIANA	New Orleans Shreveport
MASSACHUSETTS	Worcester
MICHIGAN	Ann Arbor Flint Grand Rapids Sterling Heights Warren
MISSISSIPPI	Jackson
MISSOURI	Independence Kansas City Springfield
NEBRASKA	Lincoln Omaha
NEVADA	Las Vegas Reno Clark County
NEW MEXICO	Albuquerque
NEW YORK	Various Communities
NORTH CAROLINA	Durham Raleigh Winston-Salem
OHIO	Akron Columbus Dayton Toledo
OKLAHOMA	Oklahoma City Tulsa
PENNSYLVANIA	Allentown
SOUTH CAROLINA	Greenville County Richland County Columbia
TEXAS	Abilene Amarillo Austin Beaumont Corpus Christi El Paso Laredo Pasadena San Antonio Waco
UTAH	Salt Lake County Salt Lake City
WISCONSIN	Milwaukee

Table 3 is a summary of the Phase 1 data we have collected and entered into our database as of mid December 2002. The data are separated into six major land use categories: residential, mixed residential (but mostly residential), commercial, industrial, institutional, and freeways. Our open space and other mixed land use data are not included on these tables due to lack of space in this paper. This table also summarizes all data combined. The total number of events included in the database is 3,757, with most in the residential category. Many of the monitoring locations are characterized by mixed land uses. With the exception of the mixed residential area, only the main land use categories are shown separately on this table. For most common constituents, we have detectable values for almost all monitored events. However, filtered heavy metal observations, and especially organic analyses, have many fewer detected values. This table shows the percentage of analyzed samples that had detected values. The median and coefficient of variation (COV) values are only for those data having detectable concentrations. If we included the non-detected results in these calculations, extreme biases would invalidate many of the COV calculations. Our final analyses will further examine issues associated with different detection limits, multiple laboratories, and varying analytical methods on the reported results and statistical analyses. See Burton and Pitt (2002), and the many included references in that book, for further discussions on these important issues.

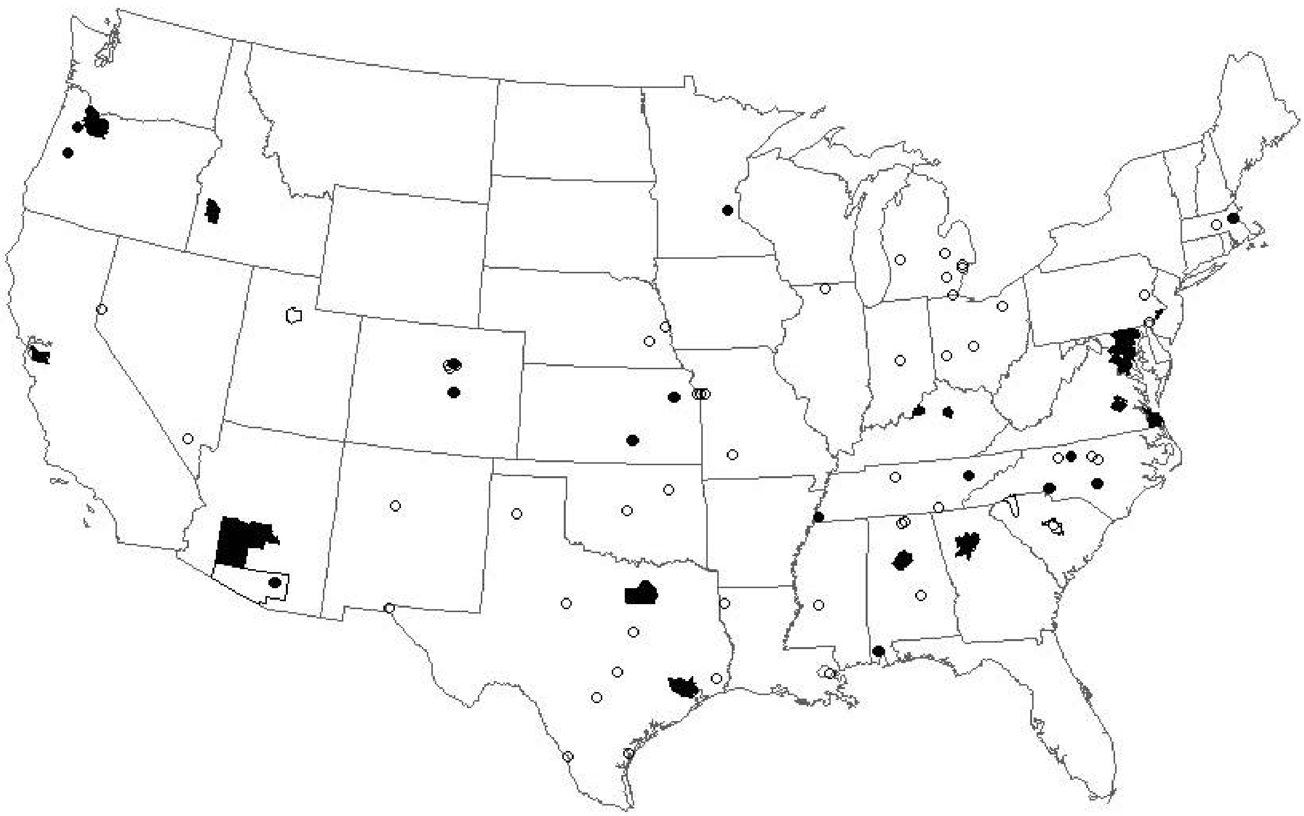

Figure 1. Data has been obtained and entered in our database for the communities shown in black. The other communities are targeted for our next data collection phase (plus Delaware, Alaska, Wisconsin, Southern California, Florida, and Hawaiian communities).

Table 3. Summary of Available Stormwater Data Included in NPDES Database

Land Use (Number of Events)	Area (acres)	% Imperv.	Precip. Depth (in)	Cond. (uS/cm @25ºC)	Hardness (mg/L $CaCO_3$)	pH
All Data Combined (3757)						
Number of observations	3562	2036	3063	887	1115	1690
% of samples above detection	94	100	100	78	81	86
Median of detected values	45	50	0.47	121	39	7.4
Coefficient of variation	7.79	0.44	0.97	1.75	1.45	0.11
Residential (983)						
Number of observations	937	558	831	164	223	247
% of samples above detection	94	100	100	65	76	74
Median of detected values	57.3	37	0.455	96	31	7.13
Coefficient of variation	4.91	0.44	0.99	1.51	0.98	0.12
Mixed Residential (584)						
Number of observations	582	239	421	137	146	341
% of samples above detection	97	100	100	77	75	88
Median of detected values	104	40	0.56	116	43.4	7.3
Coefficient of variation	2.46	0.28	0.75	1.15	0.90	0.10
Commercial (464)						
Number of observations	442	211	399	73	120	152
% of samples above detection	90	100	99	90	94	91
Median of detected values	32	80	0.39	118.5	36	7.1
Coefficient of variation	4.83	0.11	1.05	0.98	1.04	0.13
Industrial (471)						
Number of observations	448	255	395	129	114	205
% of samples above detection	93	100	100	84	79	86
Median of detected values	37.9	71.8	0.47	136	37.3	7.2
Coefficient of variation	1.70	0.32	1.00	1.31	1.09	0.11
Institutional (18)						
Number of observations	18	18	17	0	0	0
% of samples above detection	100	100	100	n/a	n/a	n/a
Median of detected values	36	45	0.18	n/a	n/a	n/a
Coefficient of variation	0.00	0.00	0.91	n/a	n/a	n/a
Freeways (185)						
Number of observations	182	154	182	86	128	111
% of samples above detection	85	100	100	100	99	100
Median of detected values	0.99	80	0.54	99	34	7.1
Coefficient of variation	0.72	0.13	1.05	1.01	1.85	0.11

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	TDS (mg/L)	TSS (mg/L)	BOD_5 (mg/L)	COD (mg/L)	Fecal Coliform (mpn/ 100 mL)	Fecal Strep. (mpn/ 100 mL)
All Data Combined (3757)						
Number of observations	3062	3525	3135	2796	1764	1142
% of samples above detection	97	98	94	96	89	91
Median of detected values	78	63	8.3	52	5000	16000
Coefficient of variation	4.13	6.05	4.45	4.79	4.64	3.85
Residential (983)						
Number of observations	802	923	867	746	382	267
% of samples above detection	97	98	96	97	87	90
Median of detected values	69	50	9.05	55.5	7750	24000
Coefficient of variation	2.17	6.25	3.34	3.49	5.06	1.89
Mixed Residential (584)						
Number of observations	470	570	557	444	342	160
% of samples above detection	98	99	92	98	93	94
Median of detected values	85	74.8	7.16	40	11000	25000
Coefficient of variation	5.68	7.89	1.37	1.47	3.21	2.21
Commercial (464)						
Number of observations	378	446	410	353	215	152
% of samples above detection	98	98	94	96	87	90
Median of detected values	74	48	12	60	3000	9200
Coefficient of variation	1.92	4.85	1.12	1.01	3.93	2.84
Industrial (471)						
Number of observations	380	434	377	339	272	176
% of samples above detection	97	98	94	96	86	92
Median of detected values	84	90	9	61	2400	13050
Coefficient of variation	4.11	4.74	6.34	2.17	6.11	6.89
Institutional (18)						
Number of observations	18	18	18	18	0	0
% of samples above detection	100	94	89	89	n/a	n/a
Median of detected values	52.5	17	8.5	50	n/a	n/a
Coefficient of variation	0.67	0.83	0.70	0.91	n/a	n/a
Freeways (185)						
Number of observations	97	134	26	67	49	25
% of samples above detection	99	99	85	99	100	100
Median of detected values	77.5	99	8	100	1700	17000
Coefficient of variation	0.80	2.53	1.26	1.06	1.95	1.21

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	$N0_2$+NO_3 (mg/L)	Ammonia (mg/L)	Nitrogen, Total Kjeldahl (mg/L)	Phos., filtered (mg/L)	Phos., total (mg/L)	Oil and Grease (mg/L)
All Data Combined (3757)						
Number of observations	3127	1874	3304	2470	3307	1830
% of samples above detection	96	75	95	89	96	71
Median of detected values	0.6	0.44	1.32	0.12	0.27	4
Coefficient of variation	1.99	3.45	3.64	2.44	8.74	4.50
Residential (983)						
Number of observations	863	564	879	656	885	473
% of samples above detection	97	87	96	90	96	66
Median of detected values	0.58	0.31	1.42	0.16	0.31	3.3
Coefficient of variation	1.93	2.14	3.87	0.98	8.13	7.79
Mixed Residential (584)						
Number of observations	542	255	562	399	554	254
% of samples above detection	96	57	94	90	95	74
Median of detected values	0.56	0.36	1.2	0.11	0.27	4
Coefficient of variation	1.01	2.96	1.85	3.70	7.98	2.53
Commercial (464)						
Number of observations	415	285	426	295	425	260
% of samples above detection	96	85	95	85	96	77
Median of detected values	0.62	0.57	1.6	0.1	0.23	5
Coefficient of variation	1.07	2.52	4.86	3.25	7.36	3.13
Industrial (471)						
Number of observations	398	243	411	301	403	287
% of samples above detection	94	91	95	90	97	74
Median of detected values	0.75	0.52	1.4	0.1	0.27	4
Coefficient of variation	0.96	3.60	2.53	1.25	6.79	3.28
Institutional (18)						
Number of observations	18	18	18	18	18	0
% of samples above detection	100	89	100	83	94	n/a
Median of detected values	0.6	0.31	1.35	0.14	0.17	n/a
Coefficient of variation	0.64	0.53	0.50	0.53	1.04	n/a
Freeways (185)						
Number of observations	25	79	125	22	128	60
% of samples above detection	96	87	97	95	99	72
Median of detected values	0.28	1.07	2	0.197	0.25	8
Coefficient of variation	1.23	1.73	1.37	2.13	1.76	0.62

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	Sb, total (mg/L)	As, total (mg/L)	As, filtered (mg/L)	Be, total (mg/L)	Cd, total (mg/L)	Cd, filtered (mg/L)	Cr, total (mg/L)
All Data Combined (3757)							
Number of observations	755	1425	209	842	2481	389	1561
% of samples above detection	9	49	27	10	49	31	63
Median of detected values	3	3.3	1.5	0.31	1	0.5	7
Coefficient of variation	2.56	2.42	1.00	2.74	4.42	1.69	1.47
Residential (983)							
Number of observations	214	366	32	239	599	85	383
% of samples above detection	2	37	6	11	38	6	50
Median of detected values	40	3	1.48	0.4	0.5	0.7	4.55
Coefficient of variation	1.11	2.42	0.50	2.92	5.20	0.55	1.31
Mixed Residential (584)							
Number of observations	74	170	18	76	398	30	172
% of samples above detection	4	65	28	16	51	40	72
Median of detected values	1	4	2	0.3	0.9	0.3	8
Coefficient of variation	1.59	3.78	0.84	2.86	3.53	0.64	1.62
Commercial (464)							
Number of observations	91	165	21	112	303	48	201
% of samples above detection	3	38	10	6	54	25	66
Median of detected values	69	2.5	1.5	0.5	0.86	0.33	6
Coefficient of variation	0.79	0.79	0.47	1.99	5.02	2.26	1.38
Industrial (471)							
Number of observations	123	219	23	164	329	42	215
% of samples above detection	18	58	13	12	60	55	72
Median of detected values	4.8	5	1	0.345	1.9	0.6	15
Coefficient of variation	1.37	0.94	0.43	2.55	3.77	1.10	1.13
Institutional (18)							
Number of observations	0	0	0	0	18	0	15
% of samples above detection	n/a	n/a	n/a	n/a	17	n/a	0
Median of detected values	n/a	n/a	n/a	n/a	0.5	n/a	n/a
Coefficient of variation	n/a	n/a	n/a	n/a	0.69	n/a	n/a
Freeways (185)							
Number of observations	14	61	72	12	95	114	76
% of samples above detection	50	56	50	17	72	26	99
Median of detected values	3	2.4	1.43	0.3	1	0.68	8.3
Coefficient of variation	0.25	0.70	1.15	0.47	0.90	1.03	0.71

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	Cr, filtered (μg/L)	Cu, total (μg/L)	Cu, filtered (μg/L)	CN, total (μg/L)	Pb, total (μg/L)	Pb, filtered (μg/L)	Hg, total (μg/L)
All Data Combined (3757)							
Number of observations	260	2770	413	1012	2902	446	1014
% of samples above detection	61	86	83	8	80	50	11
Median of detected values	2.08	16	8	5	15.9	3	0.2
Coefficient of variation	0.74	2.24	1.68	2.62	1.89	2.01	1.17
Residential (983)							
Number of observations	33	719	91	325	704	109	252
% of samples above detection	27	84	64	7	75	34	10
Median of detected values	1.28	11.1	7	5	12	3	0.2
Coefficient of variation	0.59	1.60	1.92	1.93	1.95	1.84	1.14
Mixed Residential (584)							
Number of observations	21	421	30	82	501	30	100
% of samples above detection	52	85	73	6	78	47	19
Median of detected values	2	18.7	5.75	0.01	19	3	0.3
Coefficient of variation	0.80	1.31	2.33	2.20	1.34	0.68	0.85
Commercial (464)							
Number of observations	27	360	49	144	345	59	133
% of samples above detection	41	96	80	15	95	54	11
Median of detected values	2	15	8	0.013	17	5	0.2
Coefficient of variation	0.59	1.55	1.50	1.69	1.70	1.61	0.79
Industrial (471)							
Number of observations	36	372	42	177	372	51	178
% of samples above detection	56	91	90	10	83	53	11
Median of detected values	3	21.8	8	5.92	23.7	5	0.1
Coefficient of variation	0.73	2.01	0.67	1.60	1.90	1.58	1.89
Institutional (18)							
Number of observations	0	17	0	0	0	0	0
% of samples above detection	n/a	41	n/a	n/a	n/a	n/a	n/a
Median of detected values	n/a	17	n/a	n/a	n/a	n/a	n/a
Coefficient of variation	n/a	0.59	n/a	n/a	n/a	n/a	n/a
Freeways (185)							
Number of observations	101	97	130	3	100	126	34
% of samples above detection	78	99	99	0	100	50	6
Median of detected values	2.3	34.7	10.9	n/a	27.5	1.8	0.19
Coefficient of variation	0.70	0.95	1.50	n/a	1.44	1.65	0.80

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	Ni, total (mg/L)	Ni, filtered (mg/L)	Se, total (mg/L)	Ag, total (mg/L)	Zn, total (mg/L)	Zn, filtered (mg/L)
All Data Combined (3757)						
Number of observations	1602	246	912	1149	3053	383
% of samples above detection	40	64	9	14	95	96
Median of detected values	9	4	2	3	112	51
Coefficient of variation	2.08	1.47	1.48	4.63	4.59	3.91
Residential (983)						
Number of observations	381	25	246	297	728	90
% of samples above detection	33	44	7	17	96	90
Median of detected values	6	2	2	5	73	32
Coefficient of variation	1.19	0.51	0.54	4.33	4.33	0.85
Mixed Residential (584)						
Number of observations	179	25	80	92	505	28
% of samples above detection	28	72	9	10	92	100
Median of detected values	10	5.5	4	2800	97	48
Coefficient of variation	0.84	0.87	0.89	2.02	1.06	0.88
Commercial (464)						
Number of observations	203	23	118	148	366	49
% of samples above detection	58	48	7	20	100	100
Median of detected values	7	3	2.5	5	150	59
Coefficient of variation	1.82	0.84	0.82	3.02	1.26	1.37
Industrial (471)						
Number of observations	225	36	175	216	387	42
% of samples above detection	53	58	10	23	98	95
Median of detected values	20	5	2	1	220	111.5
Coefficient of variation	0.87	1.43	0.98	4.28	2.28	3.62
Institutional (18)						
Number of observations	15	0	0	0	18	0
% of samples above detection	0	n/a	n/a	n/a	100	n/a
Median of detected values	n/a	n/a	n/a	n/a	305	n/a
Coefficient of variation	n/a	n/a	n/a	n/a	0.81	n/a
Freeways (185)						
Number of observations	79	95	16	21	93	105
% of samples above detection	87	67	6	19	97	99
Median of detected values	9.2	4	2	0.35	200	51
Coefficient of variation	0.92	1.38	n/a	0.87	1.01	1.86

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	Methylene-chloride (µg/L)	Bis(2-ethylhexyl) phthalate (µg/L)	Di-n-butyl phthalate (µg/L)	Fluoranthene (µg/L)
All Data Combined (3757)				
Number of observations	251	250	93	259
% of samples above detection	36	30	16	19
Median of detected values	11.2	9.5	0.8	6
Coefficient of variation	0.77	1.13	1.03	1.31
Residential (983)				
Number of observations	104	143	22	145
% of samples above detection	33	20	18	3
Median of detected values	11.3	4.5	10	3
Coefficient of variation	0.93	1.68	0.64	1.21
Mixed Residential (584)				
Number of observations	23	26	8	26
% of samples above detection	43	15	13	0
Median of detected values	9.05	5.1	14	n/a
Coefficient of variation	0.51	0.38	n/a	n/a
Commercial (464)				
Number of observations	42	72	20	75
% of samples above detection	21	44	25	35
Median of detected values	9.2	10.1	0.7	5.9
Coefficient of variation	0.40	1.07	1.39	4.38
Industrial (471)				
Number of observations	33	49	12	51
% of samples above detection	33	43	25	25
Median of detected values	9.7	10	0.7	3.8
Coefficient of variation	0.40	0.81	0.09	0.97
Institutional (18)				
Number of observations	0	0	0	0
% of samples above detection	n/a	n/a	n/a	n/a
Median of detected values	n/a	n/a	n/a	n/a
Coefficient of variation	n/a	n/a	n/a	n/a
Freeways (185)				
Number of observations	0	0	0	0
% of samples above detection	n/a	n/a	n/a	n/a
Median of detected values	n/a	n/a	n/a	n/a
Coefficient of variation	n/a	n/a	n/a	n/a

Table 3. Summary of Available Stormwater Data Included in NPDES Database (cont.)

	Phenanthrene (mg/L)	Pyrene (mg/L)	Diazinon (mg/L)	2, 4-D (mg/L)
All Data Combined (3757)				
Number of observations	233	249	79	101
% of samples above detection	13	14	22	35
Median of detected values	3.95	5.2	0.06	3
Coefficient of variation	1.00	1.24	1.90	0.86
Residential (983)				
Number of observations	136	140	11	11
% of samples above detection	3	4	36	64
Median of detected values	1.7	2.2	30	8
Coefficient of variation	0.70	0.30	0.40	0.72
Mixed Residential (584)				
Number of observations	23	26	1	2
% of samples above detection	0	0	0	50
Median of detected values	n/a	n/a	n/a	5
Coefficient of variation	n/a	n/a	n/a	n/a
Commercial (464)				
Number of observations	70	75	19	13
% of samples above detection	31	35	42	69
Median of detected values	4.05	5	0.045	3
Coefficient of variation	4.50	4.57	0.49	0.94
Industrial (471)				
Number of observations	47	47	9	3
% of samples above detection	17	21	33	100
Median of detected values	9	7.2	0.72	2
Coefficient of variation	0.72	0.73	1.40	1.14
Institutional (18)				
Number of observations	0	0	0	0
% of samples above detection	n/a	n/a	n/a	n/a
Median of detected values	n/a	n/a	n/a	n/a
Coefficient of variation	n/a	n/a	n/a	n/a
Freeways (185)				
Number of observations	0	0	1	1
% of samples above detection	n/a	n/a	100	0
Median of detected values	n/a	n/a	0.05	n/a
Coefficient of variation	n/a	n/a	n/a	n/a

Data Analyses

Statistical analyses are being conducted at several levels. First, probability plots are used to identify range, randomness, and normality. Figure 3 (end of paper) is an example of log-normal probability plots for some of the constituents and for all data pooled. Probability plots shown as straight lines indicate that the concentrations can be represented by log-normal distributions. This is important as it indicates that data transformations, or the use of nonparametric statistical analyses, will be needed. Other plots with obvious discontinuities (such as for bacteria, phosphorus, lead, and zinc) imply that multiple data populations may be included. Our future analyses will identify the significance of these different data categories (such as land use, region, and season).

Clustering and principal component analyses (PCA) are also being utilized to characterize expected factors influencing sample variability. Figure 4 is an example dendogram from a cluster analysis of all of the preliminary data combined. This plot indicates very close relationships between rain depth and the nutrients (total phosphorus, dissolved phosphorus, nitrite plus nitrate, ammonia, and Total Kjeldahl Nitrogen). Some of the heavy metals (cadmium, nickel, and chromium) are closely related to each other, but copper, lead and zinc are much more independent. BOD_5, COD, dissolved solids, and suspended solids are poorly related to other pollutants for the pooled data. Pearson correlation analyses did show relatively strong relationships between suspended solids and the total forms of most of the heavy metals, substantiating the observation that most of the stormwater metals are not in filtered forms.

The master data set will also be evaluated to develop descriptive statistics, such as measures of central tendency and standard errors. The runoff data will then be evaluated to determine which factors have a strong influence on event mean concentrations, including sampling methods. We will test for regional and climatic differences, the influence of land use, and the effect of storm size, among other factors. Figure 5 includes example scatter plots of COD vs. BOD_5 and filtered copper vs. total copper, illustrating these suspected close relationships. Also shown on this figure are scatter plots of suspended solids and phosphorus concentrations for different rain depths. Little variation of these concentrations with rain depth are seen when all of the data are combined, implying little likelihood of important "first-flush" effects at stormwater outfall locations. Specific comparisons of concentrations from first-flush samples with concurrent composite samples will be a more direct test and will be conducted later.

Figures 6 and 7 are example grouped box and whisker plots of suspended solids, total Kjeldahl nitrogen, fecal coliforms, and copper, grouped for different major land uses and for different seasons. The TKN and copper observations are lowest for open space areas, while the freeway locations had the highest values. Suspended solids and fecal coliform variations are not as obvious, although it is likely that the freeway bacteria values are significantly lower than those found in residential areas. The seasonal variations are not as obvious, except that the bacteria values appear to be lowest during the winter season (a similar conclusion was obtained during the NURP, EPA 1983, data evaluations). Preliminary statistical ANOVA analyses for all land use categories (using SYSTAT) found significant differences for land use categories for all pollutants. Our final analyses will further investigate this important finding and will also examine possible confounding factors.

A major goal of these analyses will be to provide guidance to stormwater managers and regulators. Especially important will be the use of this data as an updated benchmark for comparison with locally collected data. In addition, this data may be useful for preliminary calculations when using the "simple method" for predicting mass discharges for unmonitored areas. This data can also be used as guidance when designing local stormwater monitoring programs (Burton and Pitt, 2002), especially when determining the needed sampling effort based on expected variations.

We will also be examining trends of concentrations with time. A classical example would be for lead, which is expected to decrease over time with the current use of unleaded gasoline. Older stormwater samples from the 1970s typically have had lead concentrations of about 100 µg/L, or higher, while most current data indicate concentrations in the range of 1 to 10 µg/L. Figure 8 is a plot of lead concentrations for residential areas only, for the time period from 1991 to 2002. This preliminary plot shows likely decreasing lead concentrations with time for all residential sites combined. However, more work is needed to investigate interacting factors and other relationships of potential interest in order to reduce the variability inherent in this (and the other preliminary) plots.

Our final analyses will expand on these preliminary examples and will also investigate other stormwater data and sampling issues. As an example, we will compare "first flush" samples with composite samples for a number of locations and conditions (the above data only represent composite samples) and will also compare data collected manually vs. automatically.

As we are still collecting information for the database, we encourage all local and state agencies who have Phase 1 municipal stormwater data but have not previously sent it to us, to please contact us so we can arrange to have your data included in our final analyses.

References

Burton, G.A. Jr., and R. Pitt, 2002. *Stormwater Effects Handbook: A Tool Box for Watershed Managers, Scientists, and Engineers*. CRC Press, Inc., Boca Raton, FL. 911 pgs.

Smullen, J.T. and K.A. Cave, 2002. "National stormwater runoff pollution database." In: *Wet-Weather Flow in the Urban Watershed,* edited by R. Field and D. Sullivan. Lewis Publishers. Boca Raton, pgs. 67 – 78.

U.S. Environmental Protection Agency, Dec. 1983. *Results of the Nationwide Urban Runoff Program*. Water Planning Division, PB 84-185552, Washington, D.C.

Acknowledgements

Many people and institutions need to be thanked for their help on this research project. Project support and assistance from Bryan Rittenhouse, the US EPA project officer for the Office of Water, is gratefully acknowledged. The many municipalities who worked with us to submit data and information were obviously crucial and the project could not be conducted without their help. Finally, a number of graduate students at the University of Alabama (especially Veera Rao Karri, Sanju Jacob, and Sumandeep Shergill) and employees of the Center for Watershed Protection are also thanked for their careful work.

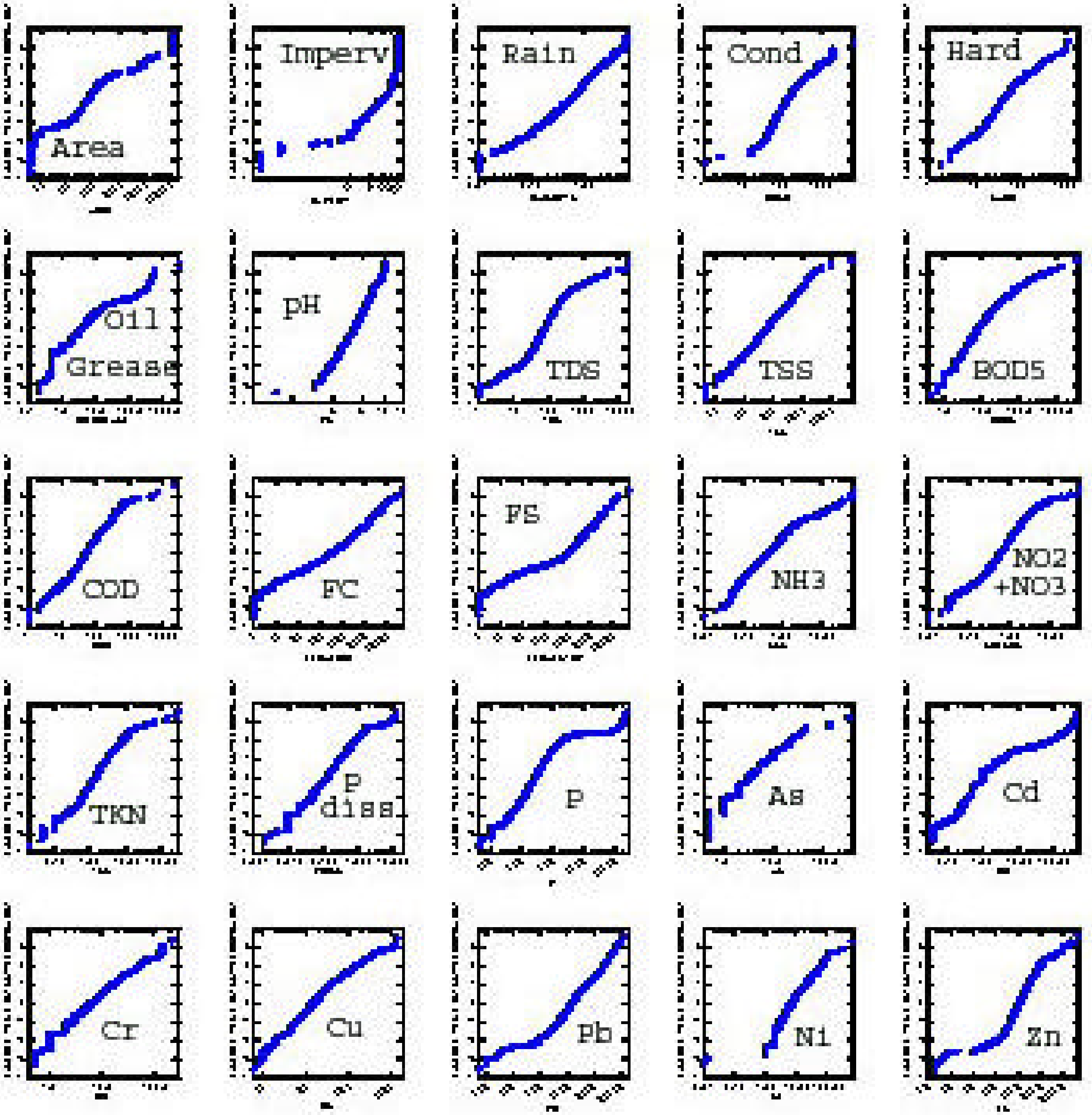

Figure 3. Log-normal probability plots of selected stormwater quality data.

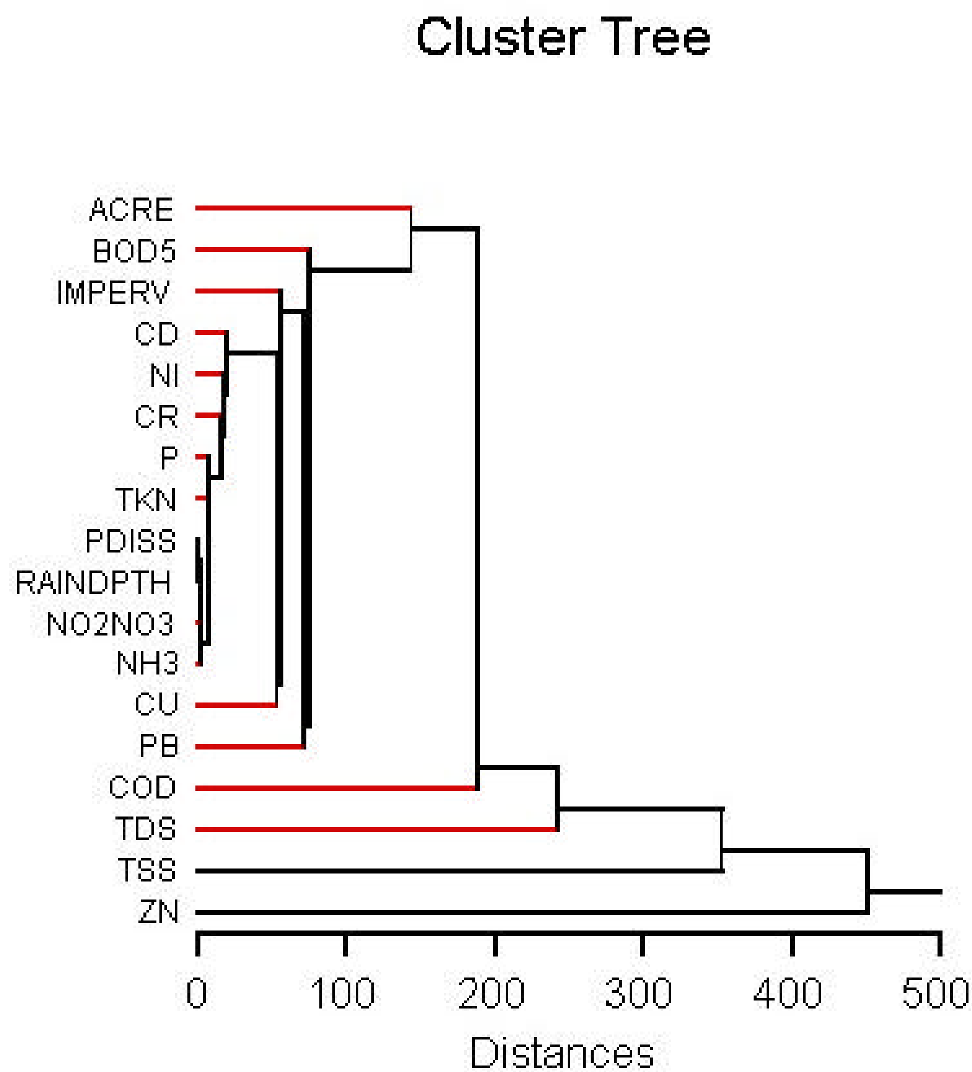

Figure 4. Cluster analysis (dendogram) showing relationships between stormwater pollutants.

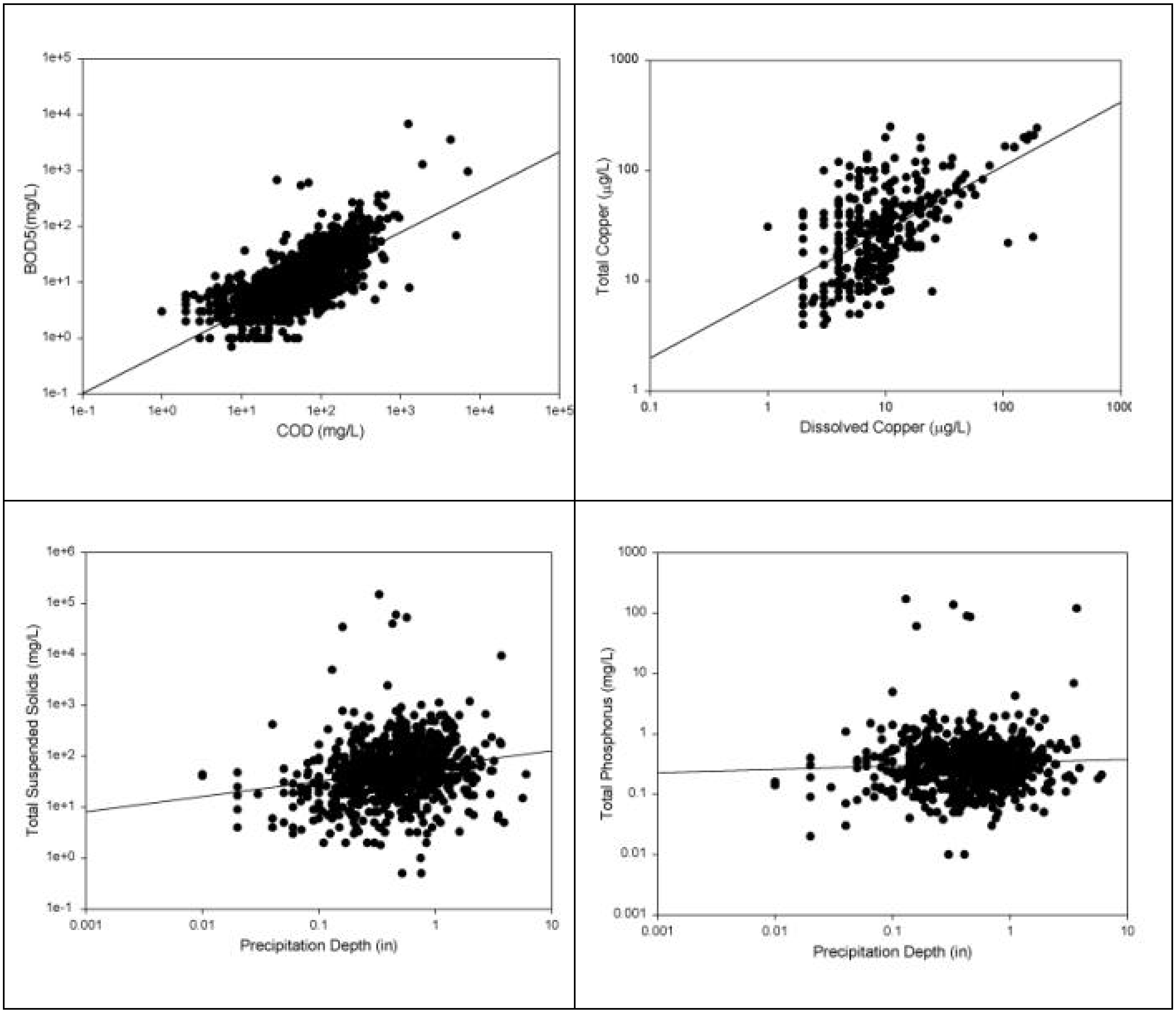

Figure 5. Example scatter plots of stormwater data.

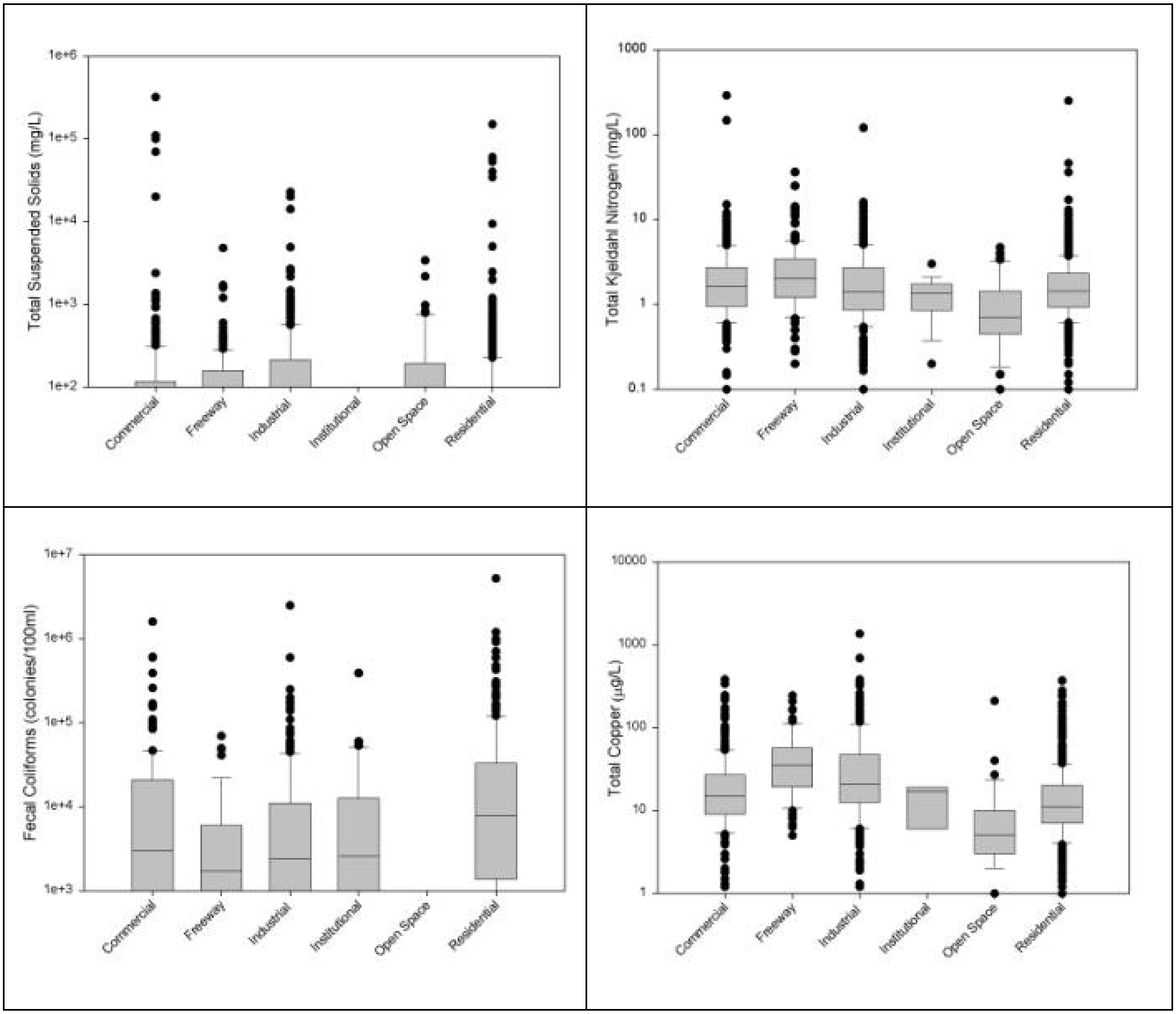

Figure 6. Example stormwater data sorted by land use (no mixed land use data included in plots).

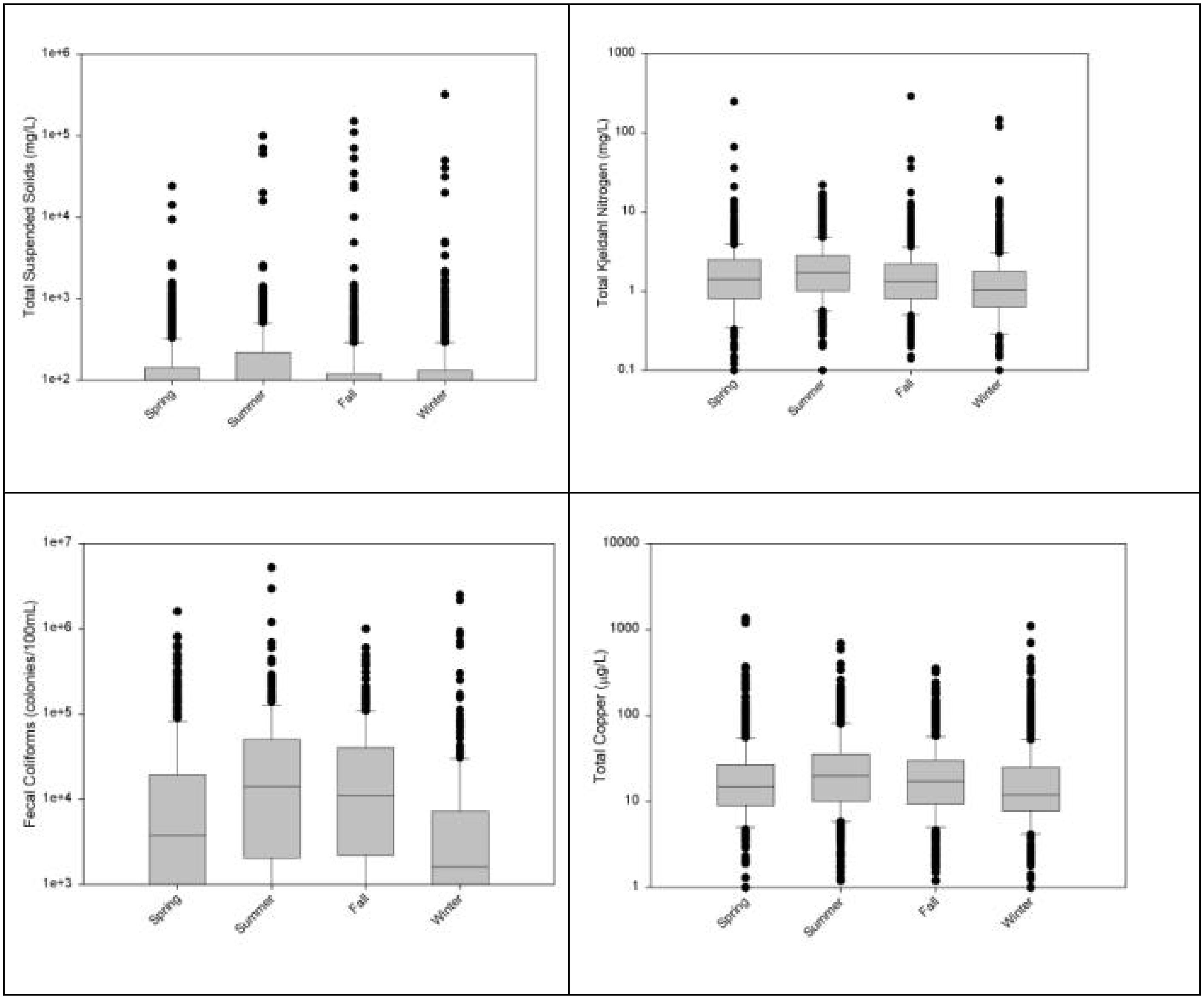

Figure 7. Example residential area stormwater pollutant concentrations sorted by season.

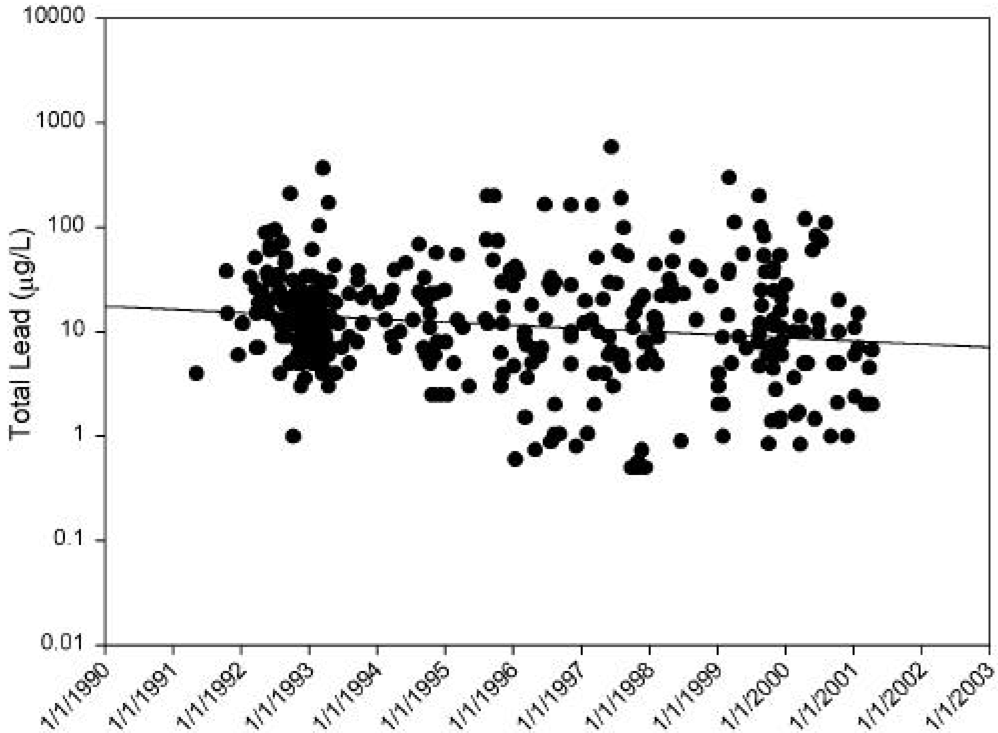

Figure 8. Residential lead concentrations with time.